

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor : **Kazuyuki SUZUKI, et al.**
Filed : **Concurrently herewith**
For : **PACKET PROCESSING APPARATUS**
Serial No. : **Concurrently herewith**

October 30, 2003


Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

PRIORITY CLAIM AND
SUBMISSION OF PRIORITY DOCUMENT

S I R:

Applicant hereby claims priority under 35 USC 119 from **Japanese** patent application number **2002-319917** filed **November 1, 2002**, a copy of which is enclosed.

Respectfully submitted,



Thomas J. Bean
Reg. No. 44,528

Katten Muchin Zavis Rosenman
575 Madison Avenue
New York, NY 10022-2585
(212) 940-8800
Docket No.: FUJI 20.710

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 2 年 1 1 月 1 日
Date of Application:

出 願 番 号 特 願 2 0 0 2 - 3 1 9 9 1 7
Application Number:
[ST. 10/C] : [J P 2 0 0 2 - 3 1 9 9 1 7]

出 願 人 富 士 通 株 式 会 社
Applicant(s):

2 0 0 3 年 7 月 3 1 日

特 許 庁 長 官
Commissioner,
Japan Patent Office

今 井 康 夫



出 証 番 号 出 証 特 2 0 0 3 - 3 0 6 1 2 8 0

【書類名】 特許願

【整理番号】 0251083

【提出日】 平成14年11月 1日

【あて先】 特許庁長官 太田 信一郎 殿

【国際特許分類】 H04B 7/00

【発明の名称】 パケット処理装置

【請求項の数】 5

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

 【氏名】 鈴木 和之

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

 【氏名】 瓜生 士郎

【特許出願人】

 【識別番号】 000005223

 【氏名又は名称】 富士通株式会社

【代理人】

 【識別番号】 100070150

 【住所又は居所】 東京都渋谷区恵比寿4丁目20番3号 恵比寿ガーデンプレイスタワー32階

 【弁理士】

 【氏名又は名称】 伊東 忠彦

 【電話番号】 03-5424-2511

【手数料の表示】

 【予納台帳番号】 002989

 【納付金額】 21,000円

【その他】 国等の委託研究の成果に係る特許出願（平成13年度通

信・放送機構「大規模な I P v 4 / I P v 6 混在ネットワークシステム構築技術の研究開発」委託研究、産業活力再生特別措置法第 3 0 条の適用を受けるもの)

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0114942

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 パケット処理装置

【特許請求の範囲】

【請求項 1】 入力されたパケットに一連番号を付加して分配する分配手段と、

前記分配手段から分配されるパケットの情報解析処理を並列に実行する複数のパケット解析手段と、

前記複数のパケット解析手段からパケットを供給され、前記パケットに付加されている一連番号順に並べ替えて出力する順序補正手段を有することを特徴とするパケット処理装置。

【請求項 2】 請求項 1 記載のパケット処理装置において、

前記分配手段は、入力されたパケット内の所定ビットの値に応じて前記複数のパケット解析手段に分配することを特徴とするパケット処理装置。

【請求項 3】 請求項 1 または 2 記載のパケット処理装置において、

前記分配手段は、前記パケット解析モジュール単位に設けた複数の出力バッファ手段を有し、

データ滞留量が閾値を超えた出力バッファ手段に対し、入力されたパケットの分配を停止することを特徴とするパケット処理装置。

【請求項 4】 請求項 1 記載のパケット処理装置において、

前記順序補正手段は、前記複数のパケット解析手段から供給されるパケットを格納するパケットバッファ手段と、

前記パケットに付加されている一連番号に対応したエントリを持つアドレス管理手段と、

前記複数のパケット解析手段から供給されたパケットを格納した前記パケットバッファ手段のアドレスを、そのパケットに付加されている一連番号に応じた前記アドレス管理手段のエントリに格納し、前記アドレス管理手段のエントリから前記一連番号順に前記パケットバッファ手段のアドレスを読み出し、前記パケットバッファ手段から前記一連番号順にパケットを読み出して出力するバッファ制御手段を

有することを特徴とするパケット処理装置。

【請求項 5】 請求項 4 記載のパケット処理装置において、

前記バッファ制御手段は、前記アドレス管理手段の全エントリに前記パケットバッファ手段のアドレスが格納されたのち、前記パケットバッファ手段のアドレスの読み出しを行うことを特徴とするパケット処理装置。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、パケット処理装置に関し、パケットの宛先、発信元等各種の情報を解析するパケット処理装置に関する。

【 0 0 0 2 】

【従来の技術】

I P (I n t e r n e t P r o t o c o l) 等のパケットを処理するルータ装置等においては、内部にパケット解析モジュールを装備し、パケットの宛先、発信元等各種の情報を解析した結果に従い、パケット転送処理を行っている。

【 0 0 0 3 】

一般的には、上記パケット解析モジュールは、ネットワークプロセッサ等のデバイスが適用されているのが現状であり、パケットの解析には C A M デバイスが用いられる。C A M (C o n t e n t A d d r e s s a b l e M e m o r y) 即ち連想メモリは、メモリ (C A M エントリ) に格納されているデータをキーとして入力し、そのキーが格納されている C A M エントリのアドレスを返却する機能を実現する。

【 0 0 0 4 】

I P ルータ装置は、電話交換機等と違ってコネクションレス通信であり、入力されるパケットをリアルタイムに解析し、その宛先、転送可否を判定するのが特徴である。

【 0 0 0 5 】

図 1 は、I P ルータ装置の一例の構成図を示す。同図中、回線終端部 1 0 1 ~ 1 0 N それぞれは回線を終端しており、各回線から受信したパケットはパケット

処理部 1 2 1 ~ 1 2 N に供給される。パケット処理部 1 2 1 ~ 1 2 N はネットワークプロセッサ (NP) を有し、パケットの宛先や発信元等各種の情報を解析してパケットと共にスイッチファブリック 1 4 に供給する。スイッチファブリック 1 4 は上記解析結果に応じてパケットをスイッチングし、スイッチングされたパケットはパケット処理部 1 2 1 ~ 1 2 N でパケットの宛先や発信元等各種の情報を解析されたのち、回線終端部 1 0 1 ~ 1 0 N を経て宛先の回線に送出される。

【 0 0 0 6 】

また、この他にも、可変長パケットを扱うシステムとしては、例えば、特許文献 1 に記載のようなものもある。

【 0 0 0 7 】

【特許文献 1】

特開 2 0 0 0 - 1 0 1 6 3 8 号公報

【 0 0 0 8 】

【発明が解決しようとする課題】

近年、データの高速化が進み、ルータ自体の容量、転送速度も上昇する傾向にあるが、この場合、パケット解析モジュール (PFE: Packet Forwarding Engine) の処理速度は転送速度に応じて高くしなければならない。

【 0 0 0 9 】

しかしながら、パケット解析モジュールとして用いられるネットワークプロセッサに代表されるようなデバイスは、処理能力に限界があり、転送速度とパケット長によっては、1 パケットについて行わなくてはならない処理が、パケット転送時間内に完了しないケースも発生してくる。更に、物理速度が高くなれば、この傾向は顕著になりパケット解析モジュールの中継性能の劣化を招く可能性が高くなるという問題点があった。

【 0 0 1 0 】

本発明は、上記の点に鑑みなされたものであり、入力パケットを複数のパケット解析モジュールで分散処理することで、全体としての中継処理性能を向上することができるパケット処理装置を提供することを目的とする。

【0011】**【課題を解決するための手段】**

請求項1に記載の発明は、入力されたパケットに一連番号を付加して分配する分配手段と、分配手段から分配されるパケットの情報解析処理を並列に実行する複数のパケット解析手段と、複数のパケット解析手段からパケットを供給され、パケットに付加されている一連番号順に並べ替えて出力する順序補正手段を有することにより、

入力パケットを複数のパケット解析手段で分散処理することで、全体としての中継処理性能を向上することができる。

【0012】

請求項2に記載の発明では、分配手段は入力されたパケット内の所定ビットの値に応じて前記複数のパケット解析手段に分配することにより、

所定ビットの値が同じパケットを同じパケット解析手段に分配することができる。

【0013】

請求項3に記載の発明では、分配手段は、パケット解析モジュール単位に設けた複数の出力バッファ手段を有し、データ滞留量が閾値を超えた出力バッファ手段に対し、入力されたパケットの分配を停止することにより、

複数のパケット解析モジュールそれぞれの負荷を平均化することができる。

【0014】

請求項4に記載の発明では、順序補正手段は、複数のパケット解析手段から供給されるパケットを格納するパケットバッファ手段と、パケットに付加されている一連番号に対応したエントリを持つアドレス管理手段と、複数のパケット解析手段から供給されたパケットを格納したパケットバッファ手段のアドレスを、そのパケットに付加されている一連番号に応じたアドレス管理手段のエントリに格納し、アドレス管理手段のエントリから一連番号順にパケットバッファ手段のアドレスを読み出し、パケットバッファ手段から一連番号順にパケットを読み出して出力するバッファ制御手段を有することにより、

複数のパケット解析手段に分配されたパケットが可変長であっても、入力順に

パケットを出力することができる。

【0015】

請求項5に記載の発明では、バッファ制御手段は、アドレス管理手段の全エンタリにパケットバッファ手段のアドレスが格納されたのち、パケットバッファ手段のアドレスの読み出しを行うことにより、

複数のパケット解析手段に分配されたパケットを入力順にパケットを出力することができる。

【0016】

【発明の実施の形態】

図2は、本発明のパケット処理装置の一実施例のブロック図を示す。このパケット処理装置は、例えば図1のパケット処理部12₁～12_Nそれぞれに使用される。

【0017】

図2において、入力パケットはディストリビュータ20に入力され、パケット解析モジュール(PFE)22₁～22_aに分配される。パケット解析モジュール(PFE)22₁～22_aそれぞれで宛先、発信元等各種の情報を解析されたパケットは順序補正バッファ24に格納され、順序補正バッファ24から入力順に読み出され、スイッチファブリックに向けて出力される。

【0018】

ここで、ディストリビュータ20の入力物理帯域をBWとしたとき、BW/a以上の入出力物理帯域を持つa個もしくはそれ以上のパケット解析モジュール22₁～22_aを並列に配置した構成である。ディストリビュータ20の出力は、パケット解析モジュール22₁～22_aの入力とインタフェースがとられ、パケット解析モジュール22₁～22_aそれぞれの出力は順序補正バッファ24の入力とインタフェースがとられている。

【0019】

更に、順序補正バッファ24の出力は物理帯域BW以上を持つ構成とする。この構成により、物理帯域BWを持つパケット処理装置を複数のパケット解析モジュールで仮想的に実現できる。

【0020】

図3は、ディストリビュータ20の一実施例のブロック図を示す。ディストリビュータ20は、SN付与部30、解析振り分け部32、出力FIFO34₁～34_aの3種のブロックから構成されており、入力される packets を後段の複数の packets 解析モジュールのいずれかに振り分けて該当出力ポートに転送する。

【0021】

SN付与部(SN_GEN)30は、最終段の順序補正制御のために、入力 packets にシーケンスナンバー(SN)を付与し、入力 packets の特定フィールドに上記のシーケンスナンバーが記入される。このシーケンスナンバーは、0～Nの値をサイクリックに付与していくが、整数Nの値については、システムで提供する最小 packets 長L_{min}、最大 packets 長L_{max}を packets 解析モジュールP個で処理する場合、次式で表される。

【0022】

$$\text{整数 } N \geq \text{整数 } N_{int} \times (P - 1) + 1$$

ただし、整数 $N_{int} > L_{max} / L_{min}$

これにより、整数Nは、ある packets 解析モジュールに最大長 packets が振り分けられ、それ以外のすべての packets 解析モジュールに全て最小 packets が振り分けられる最悪ケースにおいて、最大長 packets がある packets 解析モジュールから出力され始めた時から完了するまでに、順序補正バッファに到着しうる全 packets 数である。

【0023】

解析振り分け部32は、各 packets を複数の packets 解析モジュールに振り分けるが、この振り分けについては、いくつかの方法がある。基本的には、 packets 内の特定エリア(宛先アドレス、発信元アドレス、または、その双方)を参照し、その中の所定ビットをコード値とみなし、そのコードに対応する packets 解析モジュールに出力する。これにより、宛先アドレスや発信元アドレスが同じ packets を同一の packets 解析モジュールに分配して、宛先アドレスや発信元アドレスが同じ packets の順序を保つことができる。

【0024】

なお、このような単純なコードではなく、パケット内の所定ビットを特定の式で除算した余りによる振り分け等も可能である。

【0025】

解析振り分け部32はこの振り分け判定に従い、決定された出力ポートから出力FIFO34₁～34_aのいずれかに該当パケットを書き込み、出力FIFO34₁～34_aそれぞれの出力はパケット解析モジュール22₁～22_aの物理帯域速度で読み出される。

【0026】

ここで、上記の方法で振り分けを行った場合、入力パケットのデータパターンによっては、パケット解析モジュール22₁～22_aにデータの偏りが発生する場合があります、偏りが発生した場合には、それを補正し負荷の低いパケット解析モジュールへの迂回を制御して、各パケット解析モジュールの負荷を平均化する必要がある。このため、出力FIFO34₁～34_aそれぞれには二種類の負荷計測用のメータを配置する。

【0027】

データ容量監視メータ(DCT)35は、パケット解析モジュール単位に装備している出力FIFO34₁～34_aのデータ滞留量を監視するものである。フラグ制御部37には、この計測値に対する閾値が設定可能である。フラグ制御部37は任意の時刻における滞留量を閾値と比較し、閾値を上回っていれば負荷フラグを1にセットする。なお、閾値はヒステリシス制御が可能なように、負荷フラグのセット／リセットで個別に設けることも可能である。

【0028】

パケット数監視メータ(PCT)36は、出力FIFOの内部にあるパケット数を監視する。フラグ制御部37には、データ容量監視メータ35同様に閾値が設定可能である。フラグ制御部37は、この閾値との比較により、閾値を上回っていれば負荷フラグを1にセットする。つまり、負荷フラグは二種のメータのいずれかの閾値オーバーで1にセットされることになる。この負荷フラグの状態は、解析振り分け部32に通知される。なお、少なくともデータ容量監視メータ35を有していれば良く、パケット数監視メータ36は省略することも可能である。

。

【0029】

解析振り分け部32では、前述の論理により出力パケット解析モジュールを決定するが、このときに出力FIFO34₁～34_aからの負荷フラグを参照する。仮に、判定結果がx番目のパケット解析モジュールに対応する出力FIFOを指しており、かつ、この出力FIFOが過負荷状態にあるときには、x+1番目のパケット解析モジュールに振り分けを行う。X+1番目も過負荷状態であればx+2番目のパケット解析モジュールに振り分けを行う。

【0030】

順序補正バッファ24は、a個のパケット解析モジュール22₁～22_aからパケットを受信し、これを多重化して出力する。但し、この出力においてはディストリビュータ20の入力時と同じ順序でパケットが出力される必要がある。このため、ディストリビュータ20で付与されたシーケンスナンバーを基にパケットの順序補正を行う。

【0031】

パケットは可変長であるため、どのパケット解析モジュールからどの長さのパケットが到着するかは確定できないが、シーケンスナンバーが一巡する中での順序補正を行えば復元は可能である。

【0032】

図4は、順序補正バッファ24の第1実施例のブロック図を示す。同図中、パケット情報抽出部40₁～40_aはパケット解析モジュール22₁～22_aから供給されるパケットはパケット情報抽出部40₁～40_aから一次バッファ42₁～42_aを経て、バッファ制御部44を通してパケットバッファ46に格納される。

【0033】

パケット情報抽出部40₁～40_aはパケット解析モジュール22₁～22_aから供給されるパケットそれぞれのシーケンスナンバーとパケット長を抽出し、バッファ制御部44を通してアドレス管理テーブル48に格納する。

【0034】

パケットバッファ 46 は実際のパケットデータを供給される順に格納しておくメモリである。アドレス管理テーブル 48 は、パケットバッファ 46 に格納されているパケットの先頭アドレス及びパケット長とタグビットをパケット単位で記憶する。アドレス管理テーブル 48 は、シーケンスナンバーに対応した 0 ～ N のアドレスのエントリを持つ。

【0035】

パケット解析モジュール 221 ～ 22a からパケットが到着すると、そのパケットを到着順にパケットバッファ 46 に書き込む。次に、書き込んだパケットのシーケンスナンバーを参照し、アドレス管理テーブル 48 の上記シーケンスナンバーに該当するアドレスのエントリに、到着したパケットを書込んだパケットバッファ 46 における先頭アドレスとパケット長及び有効なエントリであることを表すタグビット（値 1）を書き込んで、このエントリを登録する。

【0036】

アドレス管理テーブル 48 の全アドレスのタグビットが値 1 となって、パケットバッファ 46 にシーケンスナンバー 0 ～ N のパケットが全て格納されていることがバッファ制御部 44 にて認識されると、バッファ制御部 44 は、読み出しポインタ（Rdp）に基づいて、アドレス管理テーブル 48 のアドレス 0 から順に各エントリの内容、つまり、先頭アドレス及びパケット長を読み出し、パケットバッファ 46 の先頭アドレスで指示されるアドレスからパケット長だけパケットデータを読み出すことにより、パケットバッファ 46 からシーケンスナンバー順にパケットを読み出して行く。

【0037】

そして、パケットバッファ 46 からの各パケットの読み出しが完了したらバッファ制御部 44 はアドレス管理テーブル 48 の読み出しパケットに対応するアドレスのタグビットを値 0 にクリアすることでエントリが空いたことを示す。そして、読み出しポインタ（Rdp）をインクリメントして、上記の処理を繰り返す。この処理により、パケットバッファ 46 から読み出されるパケットの順序は、ディストリビュータ 20 に入力時の順序と同一となる。

【0038】

この順序補正において、パケット処理装置内で正常なデータ伝送が行われている場合には、シーケンスナンバーの欠落等が無く論理的な矛盾を来さないが、パケット処理装置内の伝送路の間欠障害やビットエラー等により、シーケンスナンバーが欠落したり重複した場合の異常処理が問題となる。

【 0 0 3 9 】

このため、アドレス管理テーブル 4 8 に対して、順序補正リカバリ機構を設ける。パケットバッファ 4 6 にパケットを格納する毎に、バッファ制御部 4 4 はカウンタ 4 7 を 1 だけインクリメントし、パケットバッファ 4 6 からパケットを読み出す毎に 1 だけデクリメントする。

【 0 0 4 0 】

パケット入力数と出力数が一致する正常時には、カウンタ 4 7 は一定値を維持することになるが、パケットロス発生時は順序補正からの出力が停止するため、カウンタが増加してゆくことになる。このカウンタ値を例えば $N/2$ を閾値として監視することで障害を検出することができる。

【 0 0 4 1 】

障害再開時はディストリビュータ 2 0 の S N 生成初期値を 0 に戻し、バッファ制御部 4 4 の読み出しポインタ (R d p) の値を 0 に戻し、更に、パケットロス検出用のカウンタ 4 7 を 0 に戻すことで処理を再開する。

【 0 0 4 2 】

なお、パケットロス以外にも、順序補正バッファ 2 4 の入力部において、パケットデータに対してパリティ、 C R C 等を用いてエラー監視を行い、その結果がエラーであった場合や、パケット解析モジュール 2 2 1 ~ 2 2 a において障害を検出した場合にも同様の再開処理を行う。

【 0 0 4 3 】

図 5 は、順序補正バッファ 2 4 の第 2 実施例のブロック図を示す。同図中、図 4 と同一部分には同一符号を付す。図 5 において、パケット情報抽出部 4 0 1 ~ 4 0 a はパケット解析モジュール 2 2 1 ~ 2 2 a から供給されるパケットはパケット情報抽出部 4 0 1 ~ 4 0 a から一次バッファ 4 2 1 ~ 4 2 a を経て、バッファ制御部 5 0 を通してパケットバッファ 4 6 に格納される。

【0044】

パケット情報抽出部 40₁～40_aはパケット解析モジュール 22₁～22_aから供給されるパケットそれぞれのシーケンスナンバーとパケット長を抽出し、バッファ制御部 50を通してアドレス管理テーブル 48に格納する。

【0045】

パケットバッファ 46は実際のパケットデータを供給される順に格納しておくメモリである。アドレス管理テーブル 48は、パケットバッファ 46に格納されているパケットの先頭アドレス及びパケット長とタグビットをパケット単位で記憶する。アドレス管理テーブル 48は、シーケンスナンバーに対応した 0～Nのアドレスのエントリを持つ。

【0046】

パケット解析モジュール 22₁～22_aからパケットが到着すると、そのパケットを到着順にパケットバッファ 46に書き込む。次に、書き込んだパケットのシーケンスナンバーを参照し、アドレス管理テーブル 48の上記シーケンスナンバーに該当するアドレスのエントリに、到着したパケットを書込んだパケットバッファ 46における先頭アドレスとパケット長及び有効なエントリであることを表すタグビット（値 1）を書き込んで、このエントリを登録する。

【0047】

バッファ制御部 50は、読み出しポインタ（Rdp）に基づいて、アドレス管理テーブル 48のアドレス 0から順にテーブルの内容、つまり、先頭アドレス及びパケット長及びタグビットを読み出し、タグビットが値 1でエントリ登録が完了していることを認識すると、パケットバッファ 46からの先頭アドレスで指示されるアドレスからパケット長だけパケットデータを読み出して行く。

【0048】

この読み出しが完了したら当該タグビットの値を 0にクリアすることでアドレス管理テーブル 48のエントリが空いたことを示す。また、バッファ制御部 50は読み出しポインタをインクリメントする。この処理により、パケットバッファ 46から読み出されるパケットの順序は、ディストリビュータ 20の入力時と同一となる。

【0049】

この順序補正において、パケット処理装置内で正常なデータ伝送が行われている場合には、シーケンスナンバーの欠落等が無く論理的な矛盾を来さないが、パケット処理装置内の伝送路の間欠障害やビットエラー等により、シーケンスナンバーが欠落したり重複した場合の異常処理が問題となる。

【0050】

このため、アドレス管理テーブル48に対して、順序補正リカバリ機構を設ける。アドレス管理テーブル48に監視ウィンドウ値Wを設定し、バッファ制御部50は、アドレス管理テーブル48の読み出しポインタ(Rdp)+ウィンドウ値(W)よりも先のエントリに登録された場合に、パケットロスとみなして障害を検出する。なお、アドレス管理テーブル48はアドレスNがアドレス0に連続するリング構造として扱う。

【0051】

シーケンスナンバーの欠落や重複がない場合には、アドレス管理テーブル48のアドレス0から順にエントリが登録されると共に、登録されたエントリが順次読み出されるため、図6(A)に示すように、アドレス管理テーブル48の読み出しポインタ(Rdp)+ウィンドウ値(W)よりも先のエントリに登録されることはない。

【0052】

しかし、シーケンスナンバーの欠落や重複が生じ、図6(B)に示すように、アドレス管理テーブル48のアドレス3, 6でエントリが登録されないパケットロスが発生した場合には、アドレス管理テーブル48の読み出しポインタ(Rdp)はアドレス3で停止し、アドレス管理テーブル48の読み出しポインタ(Rdp)+ウィンドウ値(W)よりも先のエントリに登録されるため障害が検出される。

【0053】

この場合の障害再開は図6(C)に示すように、アドレス管理テーブル48のアドレス(Rdp+W)のエントリから逆順(降順)にエントリを遡って、連続して登録されているエントリの先頭(アドレス7)に読み出しポインタ(Rdp

) を移動し、読み出し処理を再開する。

【0054】

第1実施例ではシーケンスナンバーの欠落や重複により $N/2$ 個のパケットが破棄されるのに対し、本実施例では有効なエントリを無駄に廃棄することなく障害復旧処理を行うことが可能となる。

【0055】

図2に戻って説明するに、サーチエンジン26は、分散処理を行う場合の集約的な検索エンジンである。CAM27と、その関連データを格納するRAM28を備え、上位ソフトウェアからルーティングエントリやフィルタリングエントリが設定されているものである。

【0056】

パケットは負荷分散されるため、同じフローに属するパケット、即ち宛先アドレスと発信元アドレスが同じパケットが異なるパケット解析モジュールに分配される可能性があるため、複数のパケット解析モジュール22₁～22_aで同一のサーチエンジン26をアクセスする。

【0057】

このとき、複数のパケット解析モジュール22₁～22_aからのアクセスが同時に実施される可能性がある、これらの調停はサーチエンジン26が行う。調停論理は、タイムスロット割り当て方式、もしくはラウンドロビン制御により行う。この調停機能のために、サーチエンジン26の内部においては、各パケット解析モジュール22₁～22_aに対応して検索リクエストをキューイングするリクエストキューを配備しておく。

【0058】

更に、サーチエンジン26においてはカウンタを備え、CAM27の各エントリ毎にヒットした回数をカウントする。これにより、分散ルーティングされた同一フローのパケットのカウントが可能となり、集約的に統計情報を収集することが可能となる。

【0059】

図7は、本発明のパケット処理装置を適用したIPルータ装置の一実施例の構

成図を示す。同図中、回線終端部 60₁～60_Nそれぞれは例えば伝送速度 10 Gbps の光回線 (OC-192) を終端しており、各回線から受信したパケットはブレード単位で設けられたパケット処理部 62₁～62_Nに供給される。

【0060】

パケット処理部 62₁～62_Nそれぞれは、2系統のパケット処理装置を有し、これらは図2と同様に、ディストリビュータ 64、65と、処理速度が例えば 2.5 Gbps のパケット解析モジュール (NP) 66₁～66₄、67₁～67₄と、順序補正バッファ 68、69と、マルチキュー 70、71から構成されている。なお、パケット解析モジュール (PFE) 66₁～66₄、67₁～67₄それぞれには図2と同様にサーチエンジンが接続されている。

【0061】

各回線から受信したパケットはディストリビュータ 64によりパケット解析モジュール 66₁～66₄に分散されて宛先、発信元等各種の情報を解析され、順序補正バッファ 68から入力順に読み出された後、マルチキュー 70を構成する複数のキューにパケット毎にクラス (優先度) を付加してキューイングされ、優先度の高い音声パケット等を優先して取り出す如き既存のスケジューリングアルゴリズムによって各キューから取り出され、スイッチファブリック 72に供給される。

【0062】

そして、スイッチファブリック 72でスイッチングされたパケットは、ディストリビュータ 65によりパケット処理部 62₁～62_Nに供給され、パケット解析モジュール 67₁～67₄に分散されて宛先、発信元等各種の情報を解析され、順序補正バッファ 69から入力順に読み出された後、マルチキュー 71を構成する複数のキューにパケット毎にクラス (優先度) を付加してキューイングされ、既存のスケジューリングアルゴリズムによって各キューから取り出され、回線終端部 60₁～60_Nを経て宛先の回線に送出される。

【0063】

このように、本発明によれば、データの高速度が進んでも、新たに高速のパケット解析モジュールを開発することなく、既に存在するパケット解析モジュール

を用いて高速なデータ転送処理を実現することが可能となる。また、パケット解析モジュールの処理能力を軽減させることができ、多彩な処理機能の組み込みが可能となる。

【 0 0 6 4 】

更に、負荷分散により同一フローが異なるパケット解析モジュールで処理されたとしても、単一の C A M の設定によりルーティングが可能であり、かつ集約的な統計情報の収集も複雑なソフトウェア制御を必要としない。また、パケットロス等の障害発生時に、有効なパケットを無駄に廃棄せずに処理を再開することが可能となる。

【 0 0 6 5 】

なお、ディストリビュータ 2 0 が請求項記載の分配手段に対応し、パケット解析モジュール 2 2 ₁ ～ 2 2 _a がパケット解析手段に対応し、順序補正バッファ 2 4 が順序補正手段に対応し、出力 F I F O 3 4 ₁ ～ 3 4 _a が出力バッファ手段に対応し、パケットバッファ 4 6 がパケットバッファ手段に対応し、アドレス管理テーブル 4 8 がアドレス管理手段に対応し、バッファ制御部 4 4 がバッファ制御手段に対応し、カウンタ 4 7 が第 1 障害検出手段に対応し、バッファ制御部 5 0 が第 2 障害検出手段に対応し、サーチエンジン 2 6 が検索手段に対応する。

【 0 0 6 6 】

(付記 1) 入力されたパケットに一連番号を付加して分配する分配手段と

前記分配手段から分配されるパケットの情報解析処理を並列に実行する複数のパケット解析手段と、

前記複数のパケット解析手段からパケットを供給され、前記パケットに付加されている一連番号順に並べ替えて出力する順序補正手段を

有することを特徴とするパケット処理装置。

【 0 0 6 7 】

(付記 2) 付記 1 記載のパケット処理装置において、

前記分配手段は、入力されたパケット内の所定ビットの値に応じて前記複数のパケット解析手段に分配することを特徴とするパケット処理装置。

【0 0 6 8】

(付記 3) 付記 1 または 2 記載のパケット処理装置において、
前記分配手段は、前記パケット解析モジュール単位に設けた複数の出力バッファ手段を有し、

データ滞留量が閾値を超えた出力バッファ手段に対し、入力されたパケットの分配を停止することを特徴とするパケット処理装置。

【0 0 6 9】

(付記 4) 付記 1 記載のパケット処理装置において、
前記順序補正手段は、前記複数のパケット解析手段から供給されるパケットを格納するパケットバッファ手段と、

前記パケットに付加されている一連番号に対応したエントリを持つアドレス管理手段と、

前記複数のパケット解析手段から供給されたパケットを格納した前記パケットバッファ手段のアドレスを、そのパケットに付加されている一連番号に応じた前記アドレス管理手段のエントリに格納し、前記アドレス管理手段のエントリから前記一連番号順に前記パケットバッファ手段のアドレスを読み出し、前記パケットバッファ手段から前記一連番号順にパケットを読み出して出力するバッファ制御手段を

有することを特徴とするパケット処理装置。

【0 0 7 0】

(付記 5) 付記 4 記載のパケット処理装置において、
前記バッファ制御手段は、前記アドレス管理手段の全エントリに前記パケットバッファ手段のアドレスが格納されたのち、前記パケットバッファ手段のアドレスの読み出しを行うことを特徴とするパケット処理装置。

【0 0 7 1】

(付記 6) 付記 5 記載のパケット処理装置において、
前記順序補正手段は、前記パケットバッファ手段に格納されたパケット数と読み出されたパケット数の差が所定値を超えたとき障害を検出する第 1 障害検出手段を有することを特徴とするパケット処理装置。

【 0 0 7 2 】

(付記 7) 付記 4 記載の packets 処理装置において、

前記順序補正手段は、前記 packets バッファ手段のアドレスを格納した前記アドレス管理手段のエントリが、前記 packets バッファ手段のアドレスを読み出す前記アドレス管理手段のエントリに、所定の監視ウィンドウ値を加算したエントリより先となったとき、障害を検出する第 2 障害検出手段を有することを特徴とする packets 処理装置。

【 0 0 7 3 】

(付記 8) 付記 7 記載の packets 処理装置において、

前記順序補正手段は、前記障害検出手段が障害を検出したのち、前記 packets バッファ手段のアドレスを最後に格納した前記アドレス管理手段のエントリから遡り連続して前記 packets バッファ手段のアドレスが格納されている先頭のエントリから前記 packets バッファ手段のアドレスの読み出しを再開することを特徴とする packets 処理装置。

【 0 0 7 4 】

(付記 9) 付記 1 乃至 8 のいずれか記載の packets 処理装置において、

前記複数の packets 解析手段からの依頼による検索処理を集約して実行する検索手段を有することを特徴とする packets 処理装置。

【 0 0 7 5 】

(付記 1 0) 付記 9 記載の packets 処理装置において、

前記検索手段は、複数の packets 解析手段からのアクセスを調停する機能を有することを特徴とする packets 処理装置。

【 0 0 7 6 】

(付記 1 1) 付記 9 記載の packets 処理装置において、

前記検索手段は、連想メモリを用いて検索を行うことを特徴とする packets 処理装置。

【 0 0 7 7 】

(付記 1 2) 付記 1 1 記載の packets 処理装置において、

前記検索手段は、前記連想メモリの各エントリ毎にヒット回数を計数すること

を特徴とするパケット処理装置。

【 0 0 7 8 】

【発明の効果】

上述の如く、請求項 1 に記載の発明によれば、入力パケットを複数のパケット解析手段で分散処理することで、全体としての中継処理性能を向上することができる。

【 0 0 7 9 】

また、請求項 2 に記載の発明によれば、所定ビットの値が同じパケットを同じパケット解析手段に分配することができる。

【 0 0 8 0 】

また、請求項 3 に記載の発明によれば、複数のパケット解析モジュールそれぞれの負荷を平均化することができる。

【 0 0 8 1 】

また、請求項 4 に記載の発明によれば、複数のパケット解析手段に分配されたパケットが可変長であっても、入力順にパケットを出力することができる。

【 0 0 8 2 】

また、請求項 5 に記載の発明によれば、複数のパケット解析手段に分配されたパケットを入力順にパケットを出力することができる。

【図面の簡単な説明】

【図 1】

I P ルータ装置の一例の構成図である。

【図 2】

本発明のパケット処理装置の一実施例のブロック図である。

【図 3】

ディストリビュータの一実施例のブロック図である。

【図 4】

順序補正バッファの第 1 実施例のブロック図である。

【図 5】

順序補正バッファの第 2 実施例のブロック図である。

【図 6】

順序補正バッファの第 2 実施例の動作を説明するための図である。

【図 7】

本発明のパケット処理装置を適用した I P ルータ装置の一実施例の構成図である。

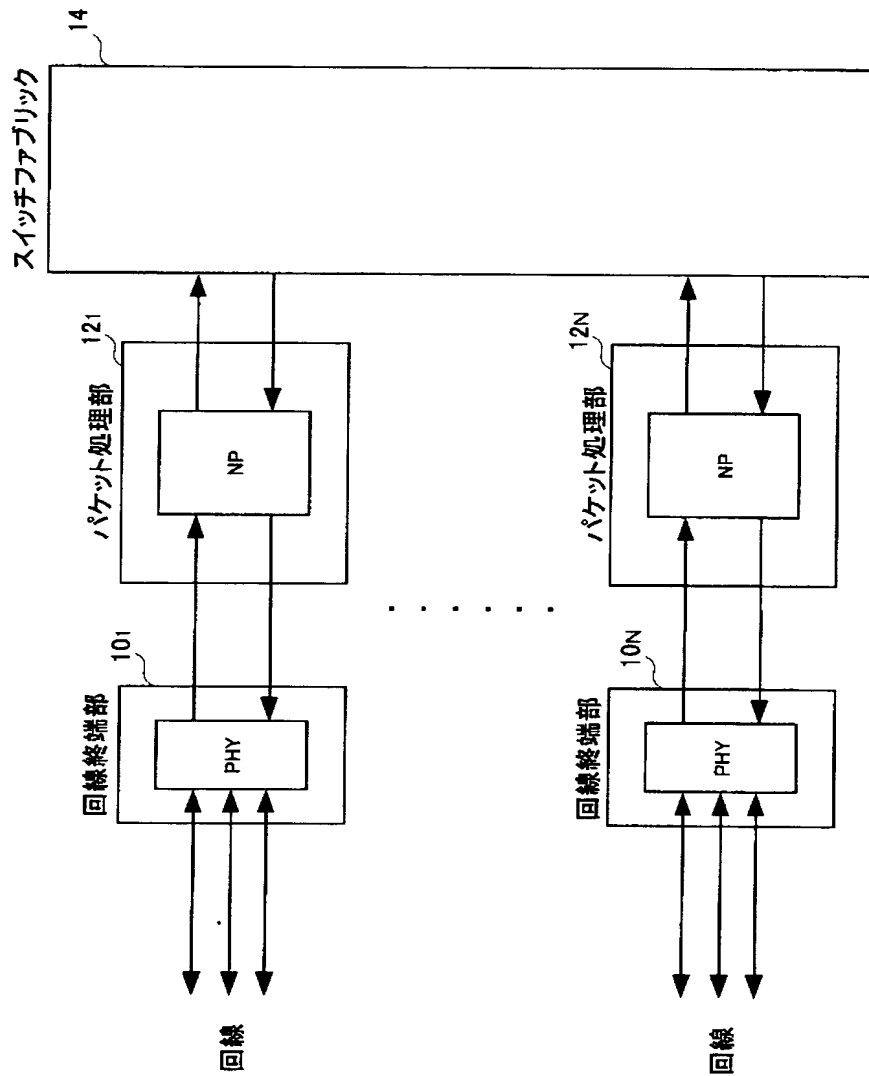
【符号の説明】

2 0, 3 0, 6 4, 6 5 ディストリビュータ
2 2 1 ~ 2 2 a, 6 6 1 ~ 6 6 4, 6 7 1 ~ 6 7 4 パケット解析モジュール
(P F E)
2 4, 6 8, 6 9 順序補正バッファ
2 6 サーチエンジン
2 7 C A M
2 8 R A M
3 0 S N 付与部 (S N _ G E N)
3 2 解析振り分け部
3 4 1 ~ 3 4 a 出力 F I F O
3 5 データ容量監視メータ (D C T)
3 6 パケット数監視メータ (P C T)
3 7 フラグ制御部
4 0 1 ~ 4 0 a パケット情報抽出部
4 2 1 ~ 4 2 a 一次バッファ
4 4, 5 0 バッファ制御部
4 6 パケットバッファ
4 7 カウンタ
4 8 アドレス管理テーブル
6 0 1 ~ 6 0 N 回線終端部
6 2 1 ~ 6 2 N パケット処理部

【書類名】 図面

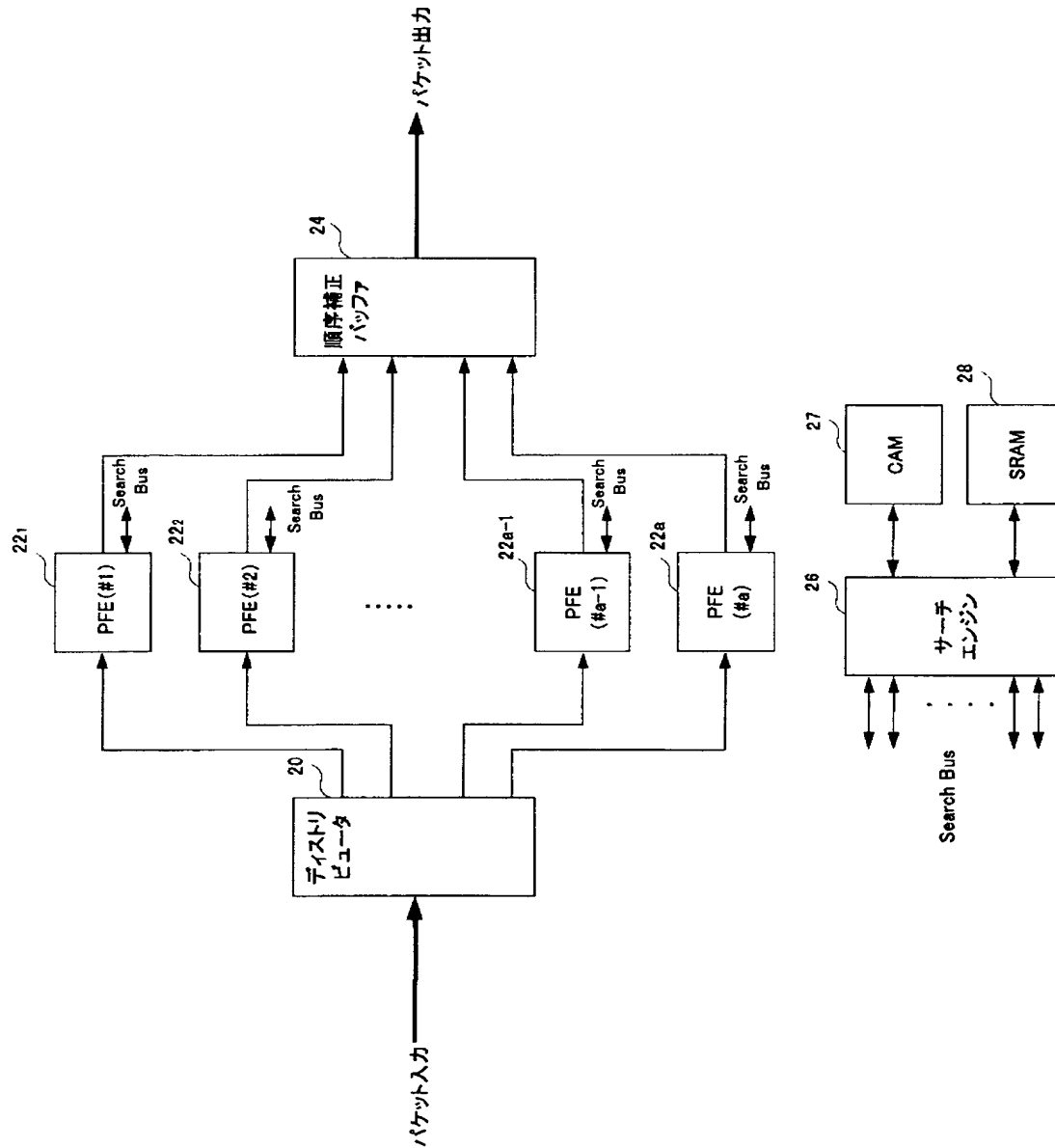
【図 1】

IPルータ装置の一例の構成図



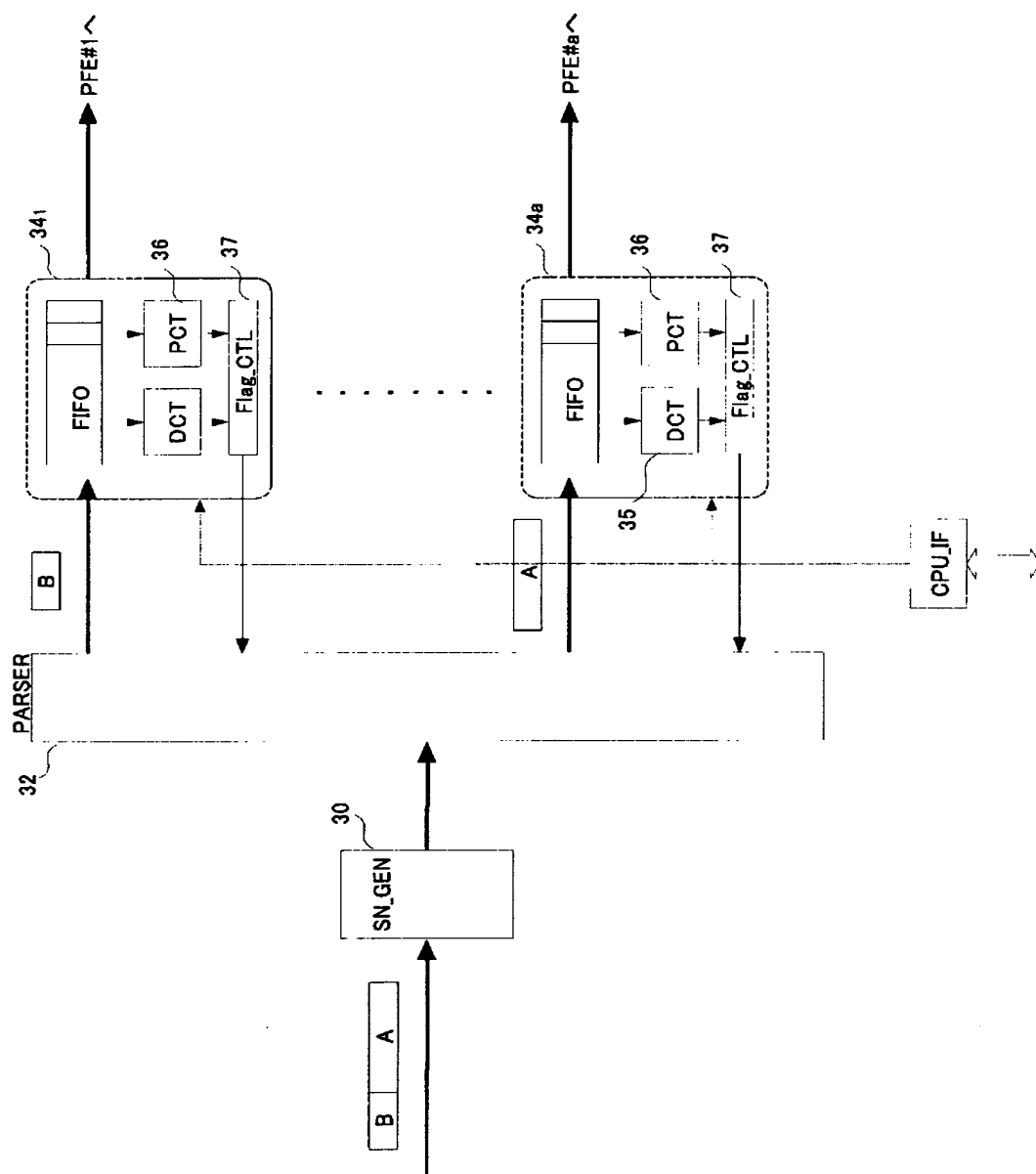
【図 2】

本発明の packets 処理装置の一実施例のブロック図



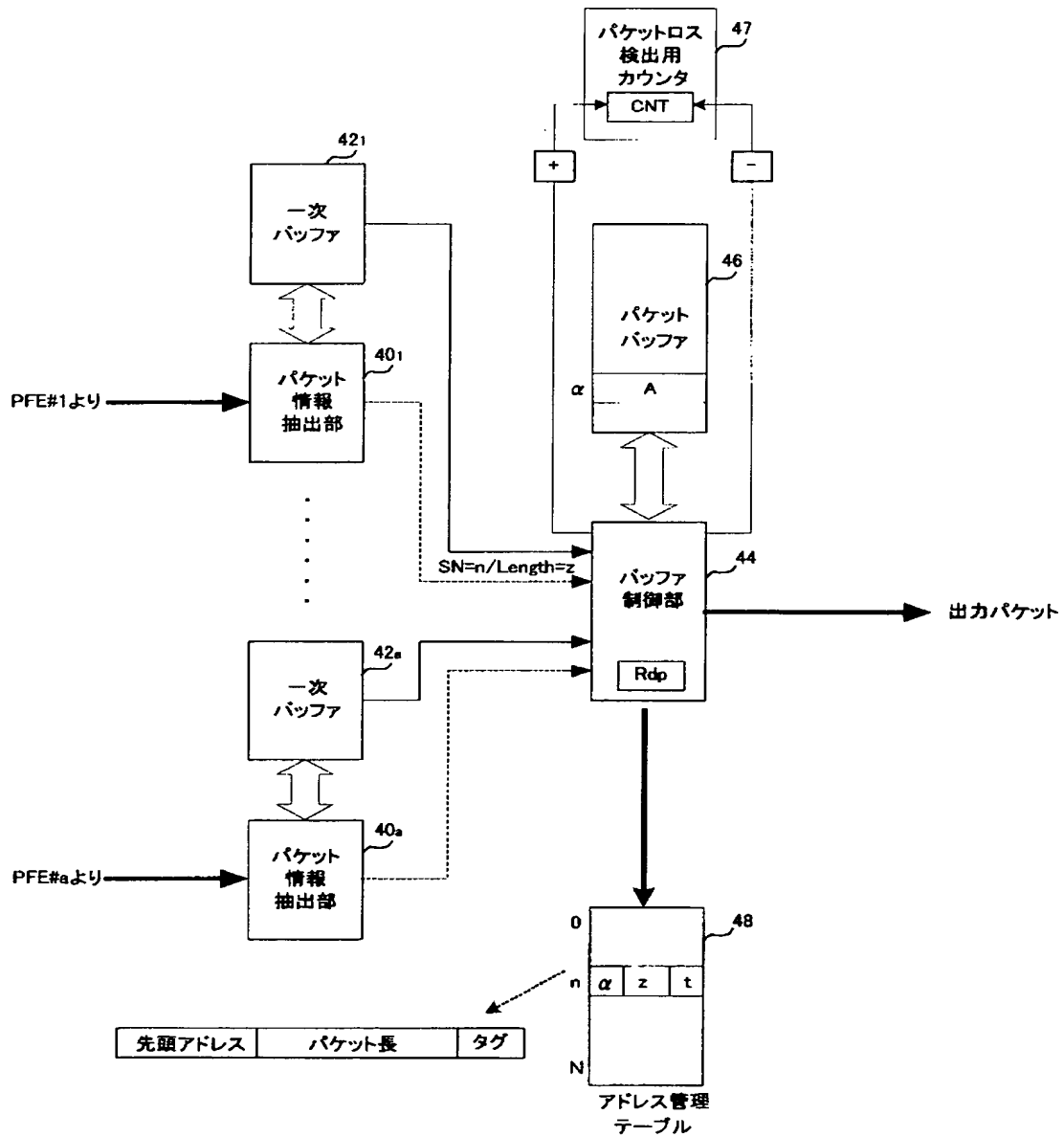
【図 3】

ディストリビュータの一実施例のブロック図



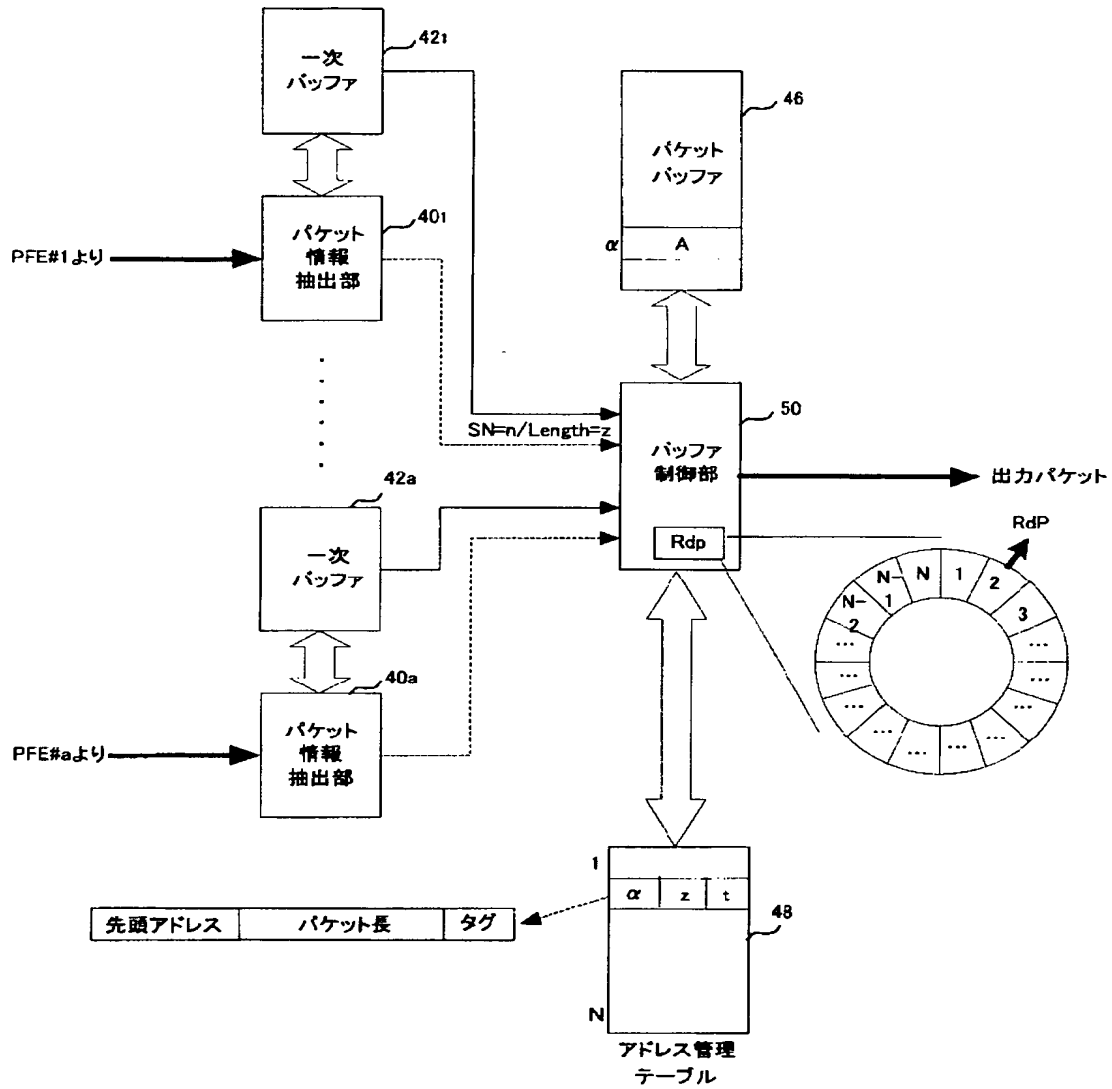
【図 4】

順序補正バッファの第1実施例のブロック図



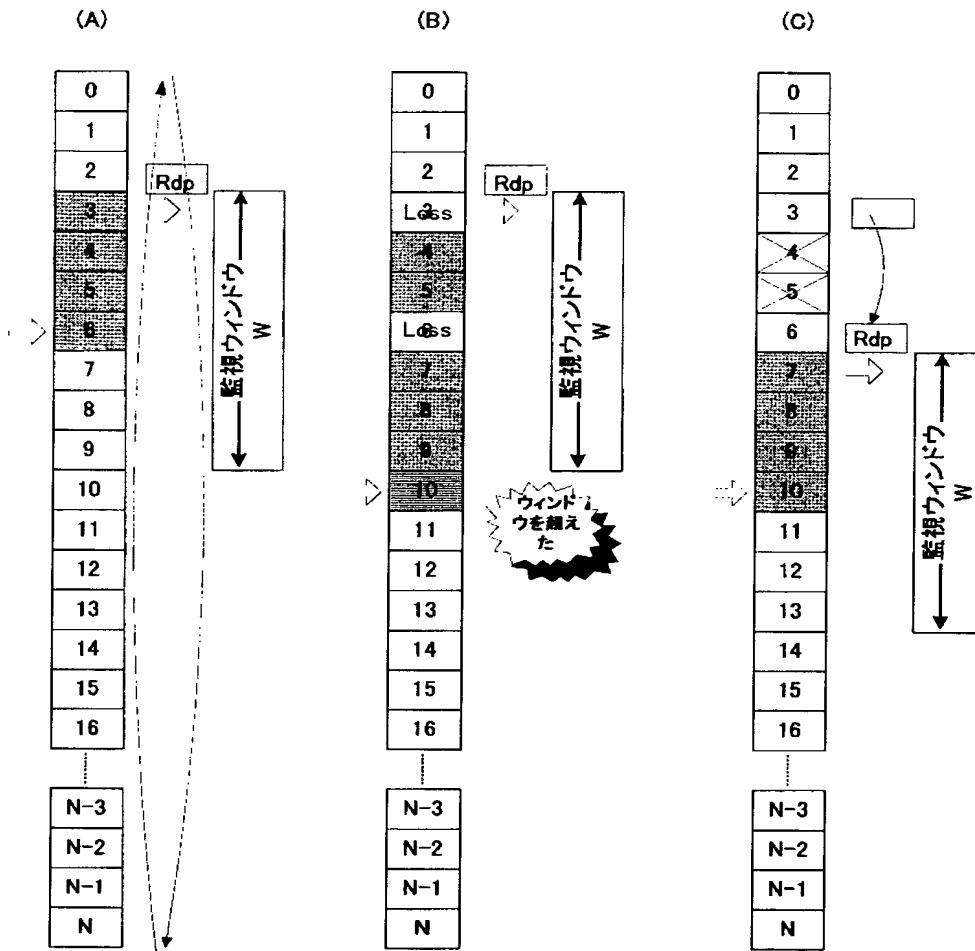
【図 5】

順序補正バッファの第2実施例のブロック図



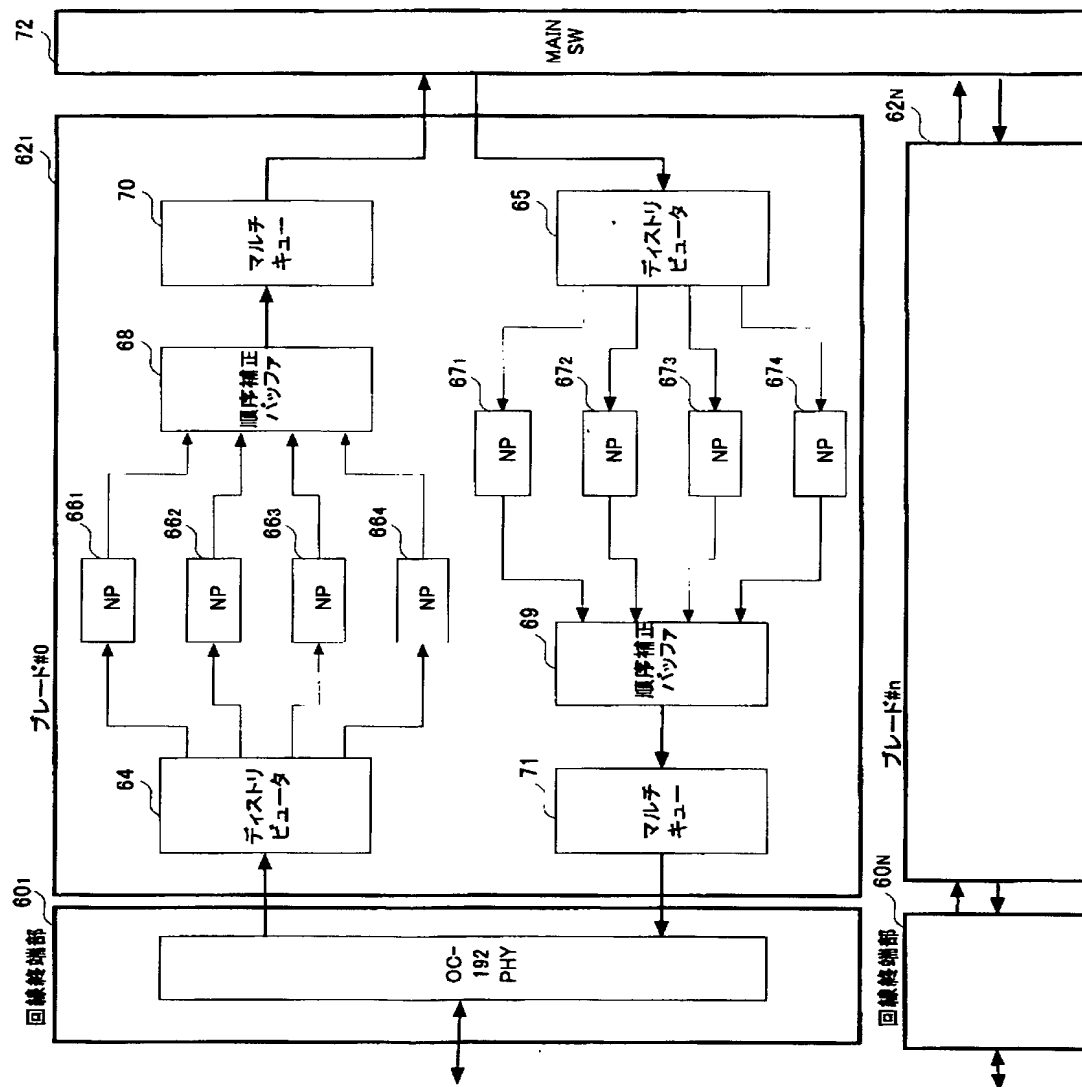
【図 6】

順序補正バッファの第3実施例のブロック図



【図7】

本発明の packets 処理装置を適用したIPルータ装置
の一実施例の構成図



【書類名】 要約書

【要約】

【課題】 本発明は、入力パケットを複数のパケット解析モジュールで分散処理することで、全体としての中継処理性能を向上することができるパケット処理装置を提供することを目的とする。

【解決手段】 入力されたパケットに一連番号を付加して分配する分配手段 20 と、分配手段から分配されるパケットの情報解析処理を並列に実行する複数のパケット解析手段 22₁～22_a と、複数のパケット解析手段からパケットを供給され、パケットに付加されている一連番号順に並べ替えて出力する順序補正手段 24 を有することにより、入力パケットを複数のパケット解析手段で分散処理することで、全体としての中継処理性能を向上することができる。

【選択図】 図 2

特願 2 0 0 2 - 3 1 9 9 1 7

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 5 2 2 3]

1. 変更年月日 1 9 9 0 年 8 月 2 4 日
[変更理由] 新規登録
住 所 神奈川県川崎市中原区上小田中 1 0 1 5 番地
氏 名 富士通株式会社
2. 変更年月日 1 9 9 6 年 3 月 2 6 日
[変更理由] 住所変更
住 所 神奈川県川崎市中原区上小田中 4 丁目 1 番 1 号
氏 名 富士通株式会社